The prognostic landscape of tumor-infiltrating immune cells in lung squamous cell carcinoma

Liang Guo[1#], Jiaoyu Yi[2#], Ming Liu[1#], Juanjuan Li[3]

[1]Department of Thoracic Surgery, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Shanghai, China; [2]Department of Thyroid and Neck Tumor, Tianjin Medical University Cancer Institute and Hospital, Tianjin, China; [3]Department of Medical Oncology, Shanghai Pulmonary Hospital & Thoracic Cancer Institute, Tongji University School of Medicine, Shanghai, China

Contributions: (I) Conception and design: L Guo, J Yi, J Li; (II) Administrative support: J Li; (III) Provision of study materials or patients: L Guo, J Li; (IV) Collection and assembly of data: L Guo, J Yi; (V) Data analysis and interpretation: L Guo, J Yi; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Juanjuan Li, MD, PhD. Department of Medical Oncology, Shanghai Pulmonary Hospital & Thoracic Cancer Institute, Tongji University School of Medicine, 507 Zheng Min Road, Shanghai 200433, China. Email: lijuanjuan700@163.com.

Background: Exploring novel biomarkers and developing effective therapeutic strategies can improve the prognosis of lung squamous cell carcinoma (LUSC) in the future. The prognostic value of tumor-infiltrating immune cells (TICs) in solid tumors has been extensively studied. However, the landscape of TICs involved in the prognosis of non-small cell lung cancer (NSCLC), especially in LUSC, remains unclear and should be systematically investigated.

Methods: This retrospective study analyzed the immune-related transcriptional profiles of 490 LUSC patients from The Cancer Genome Atlas (TCGA) cohort. Using the CIBERSORT method, TICs were evaluated and examined for their association with overall survival (OS) in LUSC.

Results: Out of the 27 TICs, 14 were correlated with prognosis in LUSC. A novel prognostic model characterized by fewer memory B cells and more central memory CD8 T cells, regulatory T cells (Tregs), and plasmacytoid dendritic cell (pDC) infiltration predicted poor OS in LUSC with high accuracy. The 1-, 3-, and 5-year areas under the curve (AUC) were 0.95, 0.98, and 0.96, respectively, in the training cohort. This finding was further validated in the validation cohort, where the 1-, 3-, and 5-year AUCs were 0.95, 0.98, and 0.95, respectively.

Conclusions: These findings may provide more effective prognostic biomarkers and potential therapeutic targets for the immunotherapy of LUSC.

Keywords: Tumor-infiltrating immune cells (TICs); prognosis; lung squamous cell carcinoma (LUSC); immune checkpoint; biomarker

Introduction

Nowadays, lung cancer has become a global public health issue and is the leading cause of cancer-related death worldwide (1). Around 95% of lung cancers are divided into two main types: small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC) (2). According to clinical, histological, genetic, and molecular aspects, NSCLC can be further classified into 2 groups, including the most common histological subtype lung adenocarcinoma (LUAD), followed by lung squamous cell carcinoma (LUSC) (2). LUSC currently has both high incidence and mortality rates, and systemic cytotoxic chemotherapy has been the main treatment strategy for advanced stage LUSC (3).

Page 2 of 11

Guo et al. The prognosis of lung squamous cell carcinoma

However, the side effects and drug resistance have caused the benefits of these treatments to reach a plateau (4). As the critical biomarkers and precise targets of the development and progression of LUSC remain unclear, molecular-targeted therapy and immunotherapy are limited (5,6). Thus, exploring novel biomarkers and developing effective therapeutic strategies can improve the prognosis of LUSC in the future.

Previously, malignant cancer phenotypes were predominantly defined by the intrinsic activities of cancer cells. In fact, it has been reported that components of the tumor microenvironment (TME) are closely related to prognosis, especially tumor-infiltrating immune cells (TICs) (7,8). Growing evidence suggests that increased infiltration of B cells, T cells, and dendritic cells (DCs) is associated with better prognosis in patients with NSCLC, however regulatory T cells (Tregs) and tumor-associated macrophages (TAMs) have the opposite effect (9). Recently, TICs have been studied in depth and multiple mechanisms have been found to contribute to lung cancer escape from immunosurveillance (10-13). In addition, immunotherapy, such as immune checkpoint blockade therapies, has recently emerged as a promising therapy strategy in lung cancer (14-16). Nevertheless, the landscape of TICs and the immune-related antigens involved in the prognosis of NSCLC, especially in LUSC, remain unclear (17-20). Therefore, it is necessary to intensively investigate whether the immune cell landscape can assist in evaluating the disease outcomes of patients with advanced-stage disease and can help develop further effective immunomodulatory strategies.

At present, emerging bioinformatics tools hold great promise for the characterization of TIC subsets. Compared with traditional techniques including immunohistochemistry and flow cytometry, bioinformatics techniques can draw more convincing and comprehensive conclusions, utilizing more available phenotypic markers and stricter criteria (21,22).

In this study, based on The Cancer Genome Atlas (TCGA) cohort, we mapped the immune cell compositions using the CIBERSORT method and quantified the immune-related antigens present in LUSC tissues. Furthermore, we characterized the prognostic landscape of 27 tumor infiltrates and identified 17 relative immune checkpoint modulators in LUSC. Importantly, a novel prognostic model was structured to predict overall survival (OS) in LUSC, which may contribute to exploring effective immunotherapies.

We present the following article in accordance with the STARD reporting checklist (available at http://dx.doi.org/10.21037/atm-21-1852).

Methods

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Database

Information of LUSC gene expression profiles and relevant normal controls were downloaded from the public database the TCGA data portal (https://tcga-data.nci.nih.gov/tcga/) in October 2019. The corresponding clinical datasets such as clinicopathological information and survival rates were also obtained from the TCGA data portal and integrated into the gene expression data. All patients were randomly assigned to 2 cohorts, including the training cohort, which comprised 40% of the total patients, and the validation cohort, which comprised 60% of the total patients. Then, we used Cell type Identification By Estimating Relative Subsets of RNA Transcripts (CIBERSORT, a deconvolution algorithm) to identify the subpopulations of TICs. In addition, 17 immune molecules previously recognized as crucial immune checkpoint modulators, including CD27, CD40, CD58, CD70, CD86, CD274, CD276, CTLA4, HAVCR2, ICOS, IDO1, LAG3, PDCD1, PDCD1LG2, TIGIT, TNFRSF14, and VTCN1, were analyzed for their correlation with tumor infiltrates and prognostic value in LUSC.

Evaluation of TICs

We evaluated the fractions of these 27 TIC subpopulations with standard gene expression data. Furthermore, the "CIBERSORT.R" package was utilized to generate a proportion matrix of TICs in LUSC, with the default signature matrix at 100 permutations (23). Only samples with a CIBERSORT P value <0.05 were enrolled for subsequent study.

Algorithm and training

As an ensemble learning method which combines multiple decision trees, a random forest classifier was used to train the data. Multiple decision trees are trained simultaneously when we utilize a random forest model, and the majority output across all trees determines the outcome of a random forest model. The maximum depth of each tree is 4 nodes.

In our study, the random forest model was trained until each tree reached the maximum depth under the constraints of the supplied parameters. Only the training set was utilized to train the model, and the generalizability of the model was tested with the validation set. The random forest model was trained on an IntelVR CoreTM i5 Central Processing Unit and implemented with the scikit-learn framework (24).

Statistical analyses

Heatmaps and clustering were generated with ClustVis, a web tool for visualizing the clustering of multivariate data. The principal component analysis (PCA) plot for the TICs in our study was generated with the plotPCA command from the DEseq2 package. The differences in TICs between tumor and normal tissue were compared with the *t*-test, and the associations between TIC levels and overall survival (OS) in LUSC patients were illustrated using Kaplan-Meier plots. Receiver operating characteristic (ROC) curves were utilized to evaluate the prognostic accuracy of the model. Hazard ratios with 95% confidence interval (CI) and log-rank P values were calculated. A P value of <0.05 was considered as the cut-off value.

Results

TIC distribution characteristics in LUSC

The data of 490 LUSC cases and 49 normal controls were available from the TCGA database in October 2019. Then, we utilized the CIBERSORT method to classify and count the 27 immune cell subtypes, including T cells [type 17 T helper cell, central memory CD4 T cell, type 1 T helper cell, effector memory CD8 T cell, Tregs, T follicular helper (Tfh) cell, central memory CD8 T cell, gamma delta T cell, activated CD4 T cell, activated CD8 T cell, effector memory CD4 T cell, and type 2 T helper cell], B cells (activated B cell, immature B cell, and memory B cell), natural killer (NK) cells (NK T cell, CD56dim NK cell, and CD56bright NK cell), and myeloid subsets [immature DC, neutrophil, activated DC, myeloid-derived suppressor cell (MDSC), macrophage, plasmacytoid dendritic cell (pDC), monocyte, eosinophil, and mast cell].

Firstly, the infiltration proportions of different TIC subtypes showed moderate to close correlations (*Figure 1*). Subsequently, PCA was utilized to detect the infiltration patterns of both the LUSC group and the normal control group. Surprisingly, there was a significant difference between the tumor and normal control groups, which might

be an intrinsic feature of LUSC (*Figure 2A*). In addition, we also compared the immune cell infiltrate profiles of LUSC patients with normal controls. As shown in *Figure 2B*, the proportions of activated CD4 T cells, memory B cells, CD56dim NK cells, and CD56bright NK cells in tumor tissue were higher than those in the normal control group. In contrast, the remaining cell types were lower in the LUSC patient group. These results indicated that the TME in LUSC was completely different from that in normal tissue, and TICs played an important role in tumor genesis and development.

Immune checkpoint modulators had no prognostic value in LUSC

As previously described, immune checkpoint modulators play important roles in all kinds of physiological and pathological processes (25). There were 17 crucial immune molecules, namely CD27, CD40, CD58, CD70, CD86, CD274 (PDL1), CD276, CTLA4, HAVCR2 (TIM3), ICOS, IDO1, LAG3, PDCD1, PDCD1LG2, TIGIT, TNFRSF14, and VTCN1, included in our analysis. As shown in *Figure 3A*, the expression of immune checkpoint modulators correlated closely with the infiltration of TICs in LUSC. Meanwhile, the prognostic value of immune checkpoint modulators was also explored, however, none of them were correlated with the prognosis of LUSC (*Figure 3B,C,D,E,F,G* and Figure S1).

The prognostic value of TICs in LUSC patients

To examine the predictive value of TICs for OS, Kaplan-Meier survival analyses were performed according to different TIC distribution patterns. The results showed that 14 of the 27 TICs were associated with the outcome in LUSC. Only 2 infiltrates (memory B cells and CD56bright NK cells) were related to improved prognosis, while 12 other infiltrates, including central memory CD4 T cells, central memory CD8 T cells, Tregs, gamma delta T cells, CD56dim NK cells, macrophages, pDCs, MDSCs, neutrophils, type 1 T helper cells, mast cells, and eosinophils, were related to poor survival in LUSC (*Figure 4* and Figure S2).

High efficiency immune prognostic model in LUSC

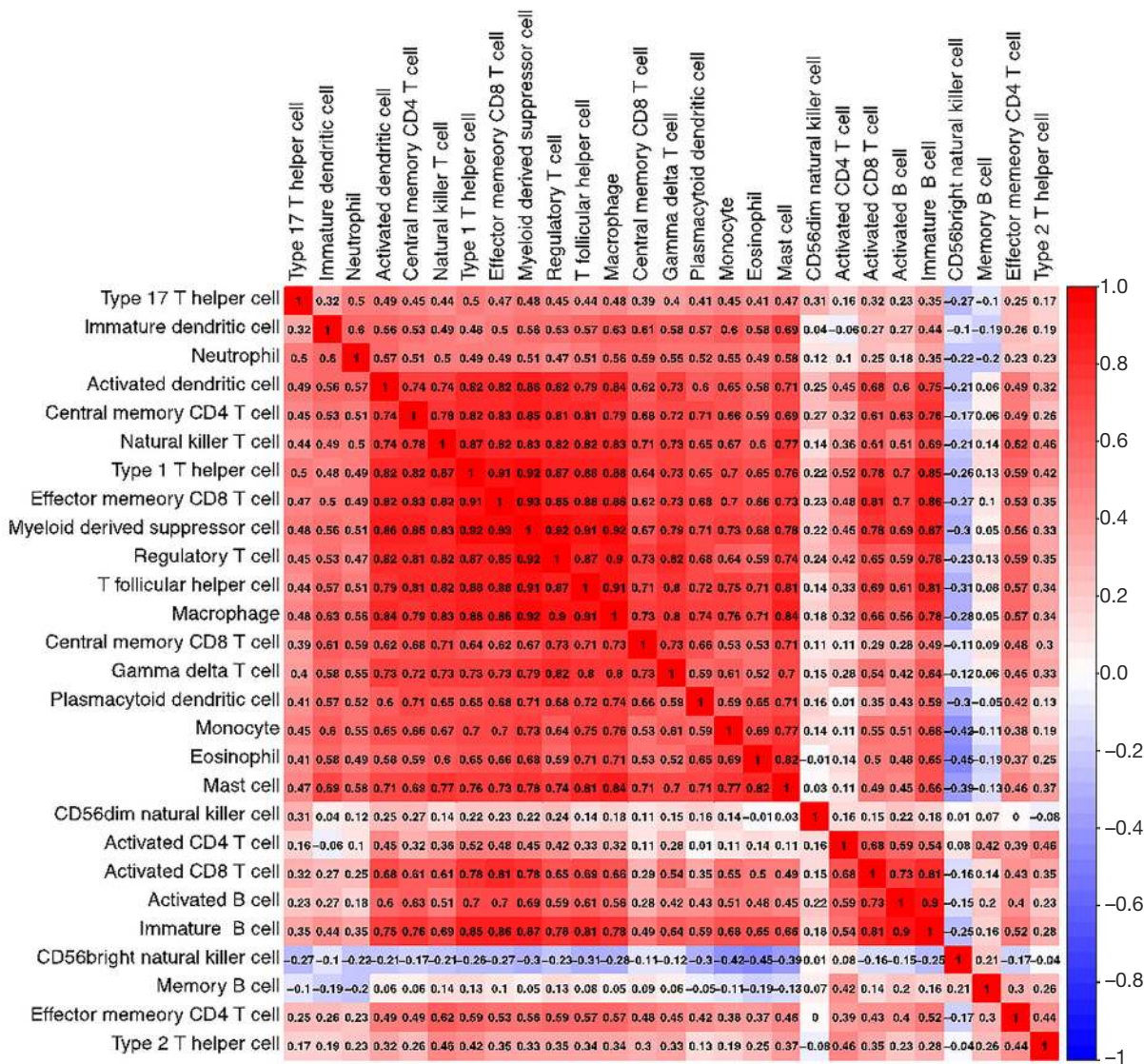Based on the above results, a novel predictive model with the ability to calculate the prognostic risk score for LUSC

Page 4 of 11

Guo et al. The prognosis of lung squamous cell carcinoma

Figure 1 Correlations of tumor-infiltrating immune cell (TIC) subsets in lung squamous cell carcinoma. The proportions of different TIC subpopulations were moderately to closely correlated.

was established after applying the random forest model to the training set. LUSC patients were divided into low-risk and high-risk groups according to the prognostic risk score analyzed with "Cutoff Finder", an online tool to identify the best cut-off (26) (*Figure 5A*). The risk curve and survival status showed that patients considered as high risk tended to have a worse prognosis than patients with low risk (*Figure 5B*). Central memory CD8+ T cells, pDCs, and Tregs were found to be enriched in the high-risk group, while memory B cells were predominantly enriched

in the low-risk group (*Figure 5C*). It was clear that the new immune-related predictive model was significantly associated with OS in LUSC patients (P<0.0001) (*Figure 5D*). The 1-, 3-, and 5-year areas under the curve (AUC) of the survival ROC curves reached 0.95, 0.98, and 0.96 respectively, supporting the accuracy and effectiveness of the new predictive model (*Figure 5E*).

To further verify the predictive ability of the new immune-related model, a similar analysis was performed in the validation dataset. Consistent with the training
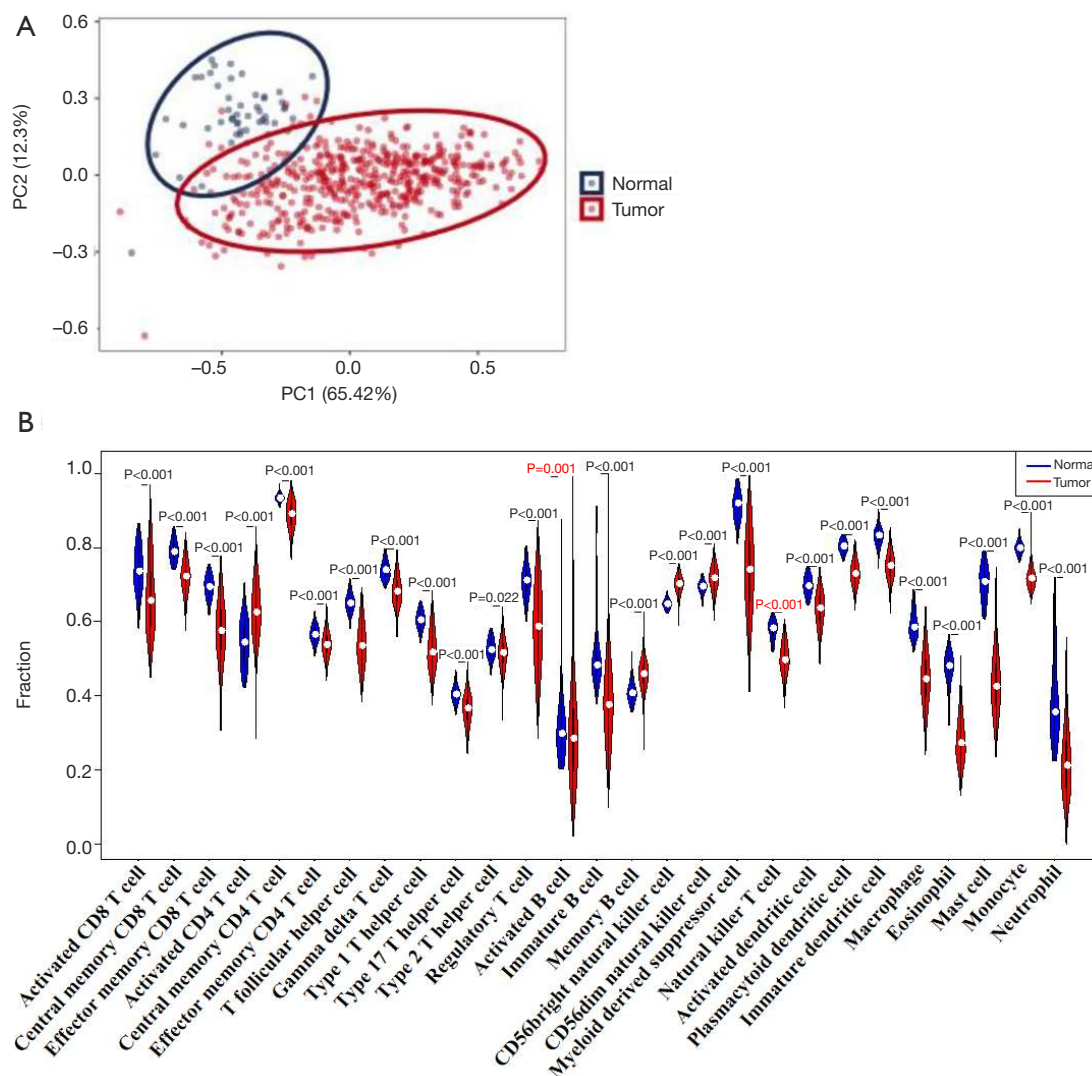
Figure 2 Distribution differences of tumor-infiltrating immune cells (TICs) between normal and lung squamous cell carcinoma (LUSC) patients. (A) The infiltration patterns of the LUSC and normal groups were characterized by a principal component analysis (PCA) model. (B) Comparison of TIC subsets between tumor and normal tissue.

set, patients considered as high risk tended to have worse outcomes than patients with low risk (*Figure 6A,B,C*). The immune-related predictive model was also significantly correlated with OS in the validation cohort (P<0.0001) (*Figure 6D*). The 1-, 3-, and 5-year AUCs of the survival ROC curves reached 0.95, 0.98, and 0.95 respectively (*Figure 6E*), indicating satisfactory predictive stratification by the final immune-related predictive model.

Discussion

In this study, we systematically elaborated on TIC

composition, immune checkpoint modulator distribution, and their prognostic value in LUSC. More importantly, based on TICs, we structured a novel immune-related prognostic model for LUSC which indicated that patients with fewer memory B cells and more central memory CD8 T cells, Tregs, and pDC infiltration had poor prognosis. This predictive model can identify LUSC patients with worse prognoses more accurately and effectively.

It was shown that Tfh cells were positively associated with better prognosis in a similar study (27), while an increased number of neutrophils indicated a poor outcome in LUSC. Consistent with the above study, we demonstrated

Page 6 of 11

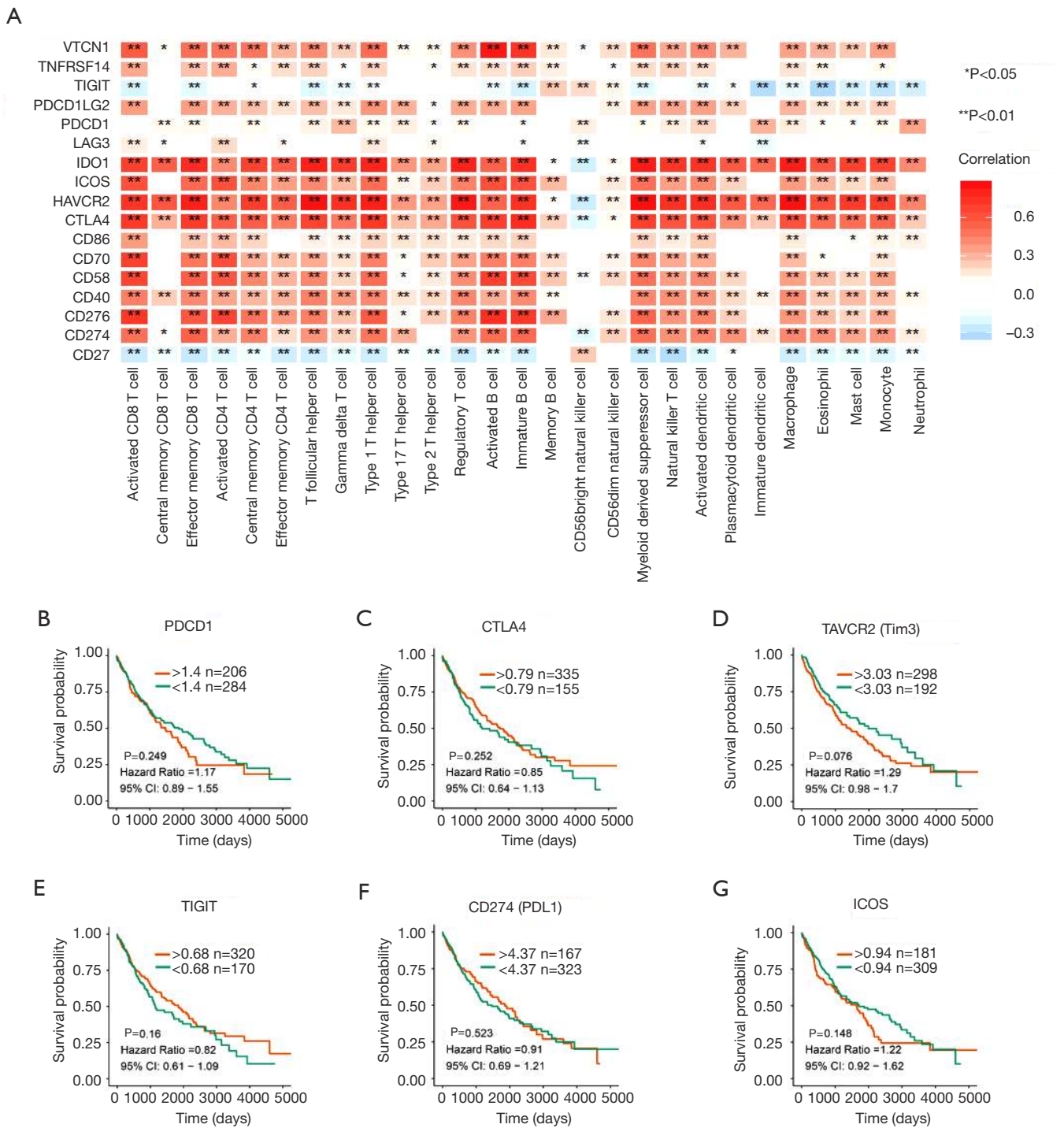Guo et al. The prognosis of lung squamous cell carcinoma

Figure 3 Prognostic analysis of immune checkpoint modulators in lung squamous cell carcinoma (LUSC). (A) Correlations between immunomodulators and the tumor-infiltrating immune cell (TIC) distribution patterns; (B,C,D,E,F,G) Kaplan-Meier curves of different immune modulators in LUSC. *P<0.5, **P<0.01.
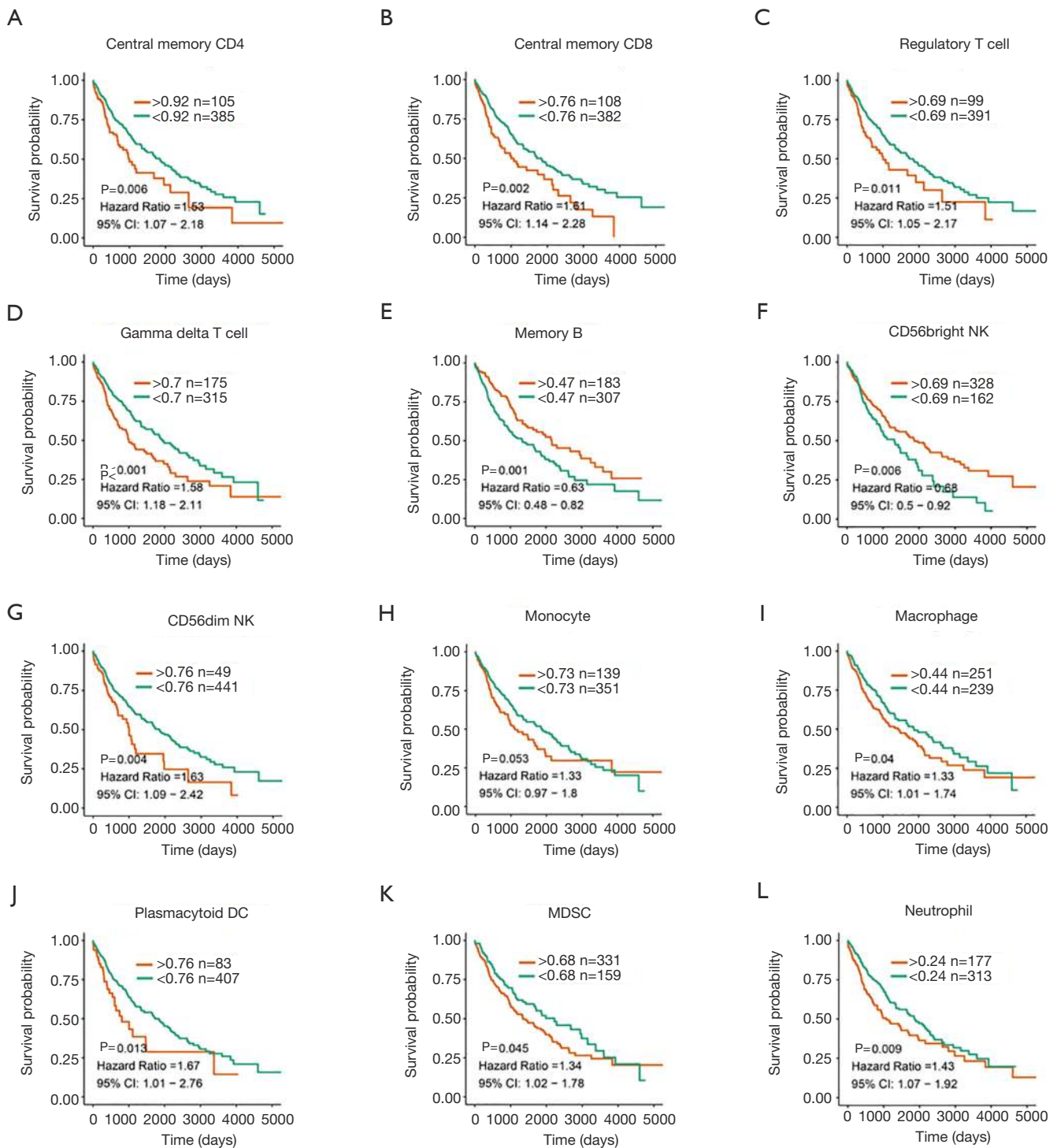
Figure 4 Kaplan-Meier curves of tumor-infiltrating immune cell (TIC) distribution patterns in lung squamous cell carcinoma (LUSC). (A-L) Kaplan-Meier curves of overall survival split within LUSC. Hazard ratios with 95% CI and log rank P values were calculated. P<0.05 was considered significant.
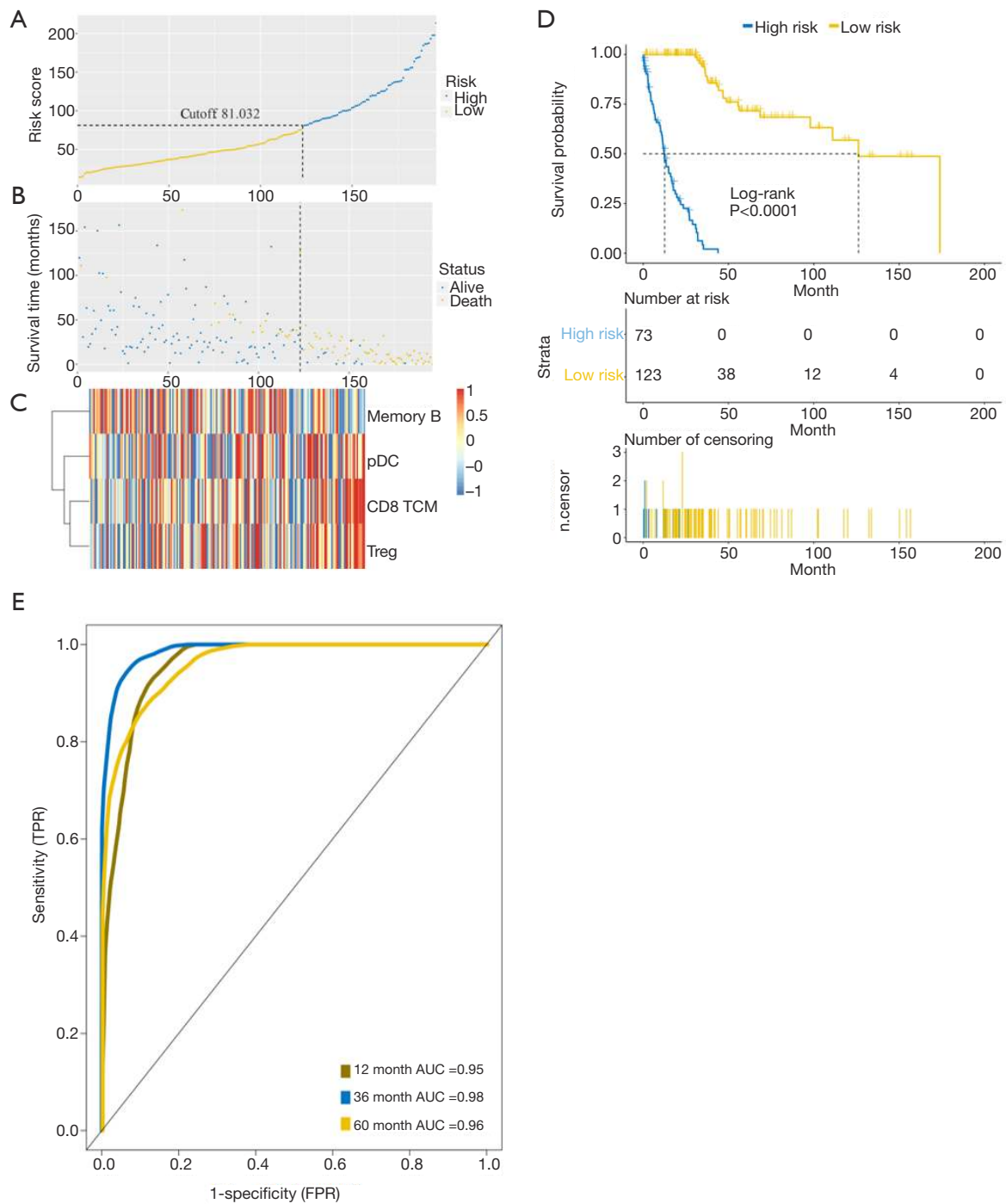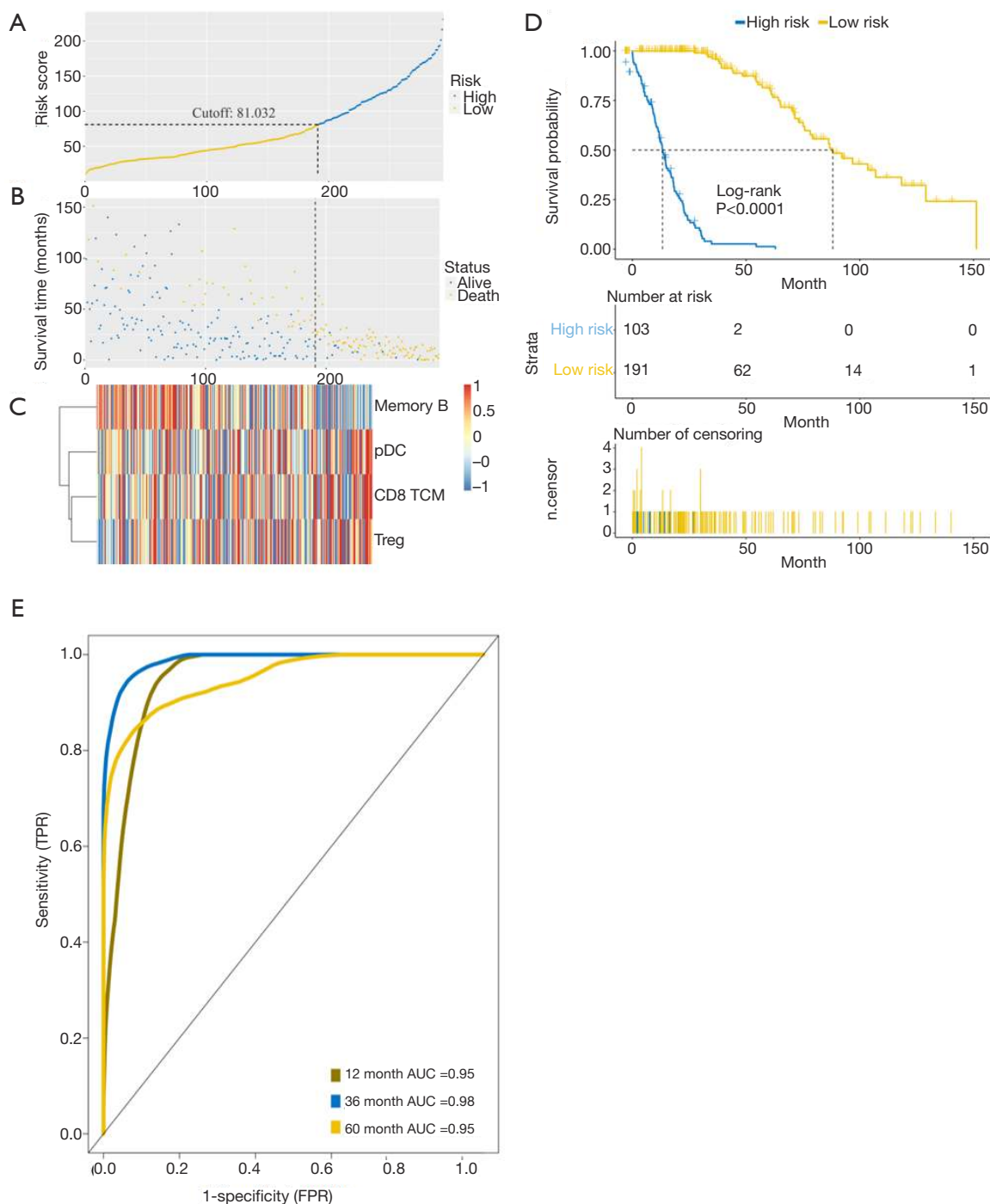
Figure 5 Risk score analysis of the immune-related predictive model in lung squamous cell carcinoma (LUSC) in the training cohort. (A) Low- and high-risk groups of LUSC patients based on the predictive model. (B) Survival status and duration of LUSC patients. (C) Proportional heatmap of the 4 tumor-infiltrating immune cells (TICs) in LUSC. The transition from blue to red indicates the trend from a low proportion to a high proportion. (D) The immune-related predictive model significantly predicts overall survival (OS) in LUSC. (E) Receiver operating characteristic (ROC) curve analysis of survival prediction by the predictive model.

Figure 6 Risk score analysis of the immune-related predictive model in lung squamous cell carcinoma (LUSC) in the validation cohort. (A) Low- and high-risk groups of LUSC patients based on the predictive model. (B) Survival status and duration of LUSC patients. (C) Proportional heatmap of the 4 tumor-infiltrating immune cells (TICs) in LUSC. The transition from blue to red indicates the trend from a low proportion to a high proportion. (D) The immune-related predictive model significantly predicts overall survival (OS) in LUSC. (E) Receiver operating characteristic (ROC) curve analysis of survival prediction by the predictive model.

Page 10 of 11

Guo et al. The prognosis of lung squamous cell carcinoma

a negative correlation between neutrophil count and patient survival. It is well established that inflammation is one of the hallmarks of tumors (28). The activity of other immune cells could be inhibited by neutrophils, leading to tumorigenesis and immune escape. While the prognostic value of Tfh cells in LUSC remains controversial, our results suggested that Tfh cells were inversely correlated with patient survival. A previous study (29) demonstrated that a high number of Tfh cells could induce immunosuppression and lead to tumor development in NSCLC, indicating the tumor-promoting role of Tfh cells in lung cancer. More importantly, the AUC of the immune-related prognostic model reached above 0.95 in both the training and validation cohorts, indicating that our novel model has more stability and efficiency.

A meta-analysis (30) showed that FoxP3+ Tregs had dual prognostic value depending on the tumor types. It was also demonstrated that Tregs and pDCs predicted worse prognosis in the TME of gastric cancer (GC) as the main immunosuppressive cells (31). As the key component of humoral immunity, B cells play an important role in the prognostic impact in human solid tumors. A previous study (32) proved that CD27-isotype-switched memory B cells served as independent favorable predictors of prognosis in human hepatocellular carcinoma (HCC). Consistent with the previous studies above, we also indicated that Tregs, pDCs, and memory B cells had the same prognostic value in LUSC. Usually, central memory CD8+ T cells are considered as anti-tumor factors (33). Although this was a negative prognostic marker for LUSC in this study, it is not contradictory. We speculated that the higher the degree of malignancy of the tumor, the more the infiltration of central memory CD8+ T cells with anti-tumor function, but the increase in the degree of immune cell infiltration is not as rapid as the rate of tumor progression. Thus, a high level of immune cell infiltration does not necessarily represent a good prognosis.

In conclusion, we provided a comprehensive analysis of the effects of TICs and immune checkpoint modulators on the prognosis of LUSC, and established a novel immune-related prognostic model based on the TICs which could effectively identify the LUSC patients with poor prognosis. These findings provide effective strategies for improving immunotherapy and guiding personalized therapy in LUSC.

Acknowledgments

Footnote

References

1. Alberg AJ, Brock MV, Ford JG, et al. Epidemiology of lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest 2013;143:e1S-e29S.
2. Nasim F, Sabath BF, Eapen GA. Lung Cancer. Med Clin North Am 2019;103:463-73.
3. Lemjabbar-Alaoui H, Hassan OU, Yang YW, et al. Lung cancer: Biology and treatment options. Biochim Biophys Acta 2015;1856:189-210.
4. Jordan EJ, Kim HR, Arcila ME, et al. Prospective Comprehensive Molecular Characterization of Lung Adenocarcinomas for Efficient Patient Matching to Approved and Emerging Therapies. Cancer Discov 2017;7:596-609.
5. Kelly RL, Le D, Zhao J, et al. Reduction of Nonspecificity Motifs in Synthetic Antibody Libraries. J Mol Biol 2018;430:119-30.
6. Brahmer J, Reckamp KL, Baas P, et al. Nivolumab versus

Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. N Engl J Med 2015;373:123-35.

7. Diakos CI, Charles KA, McMillan DC, et al. Cancer-related inflammation and treatment effectiveness. Lancet Oncol 2014;15:e493-503.

8. Candido J, Hagemann T. Cancer-related inflammation. J Clin Immunol 2013;33 Suppl 1:S79-84.

9. Dieu-Nosjean MC, Antoine M, Danel C, et al. Long-term survival for patients with non-small-cell lung cancer with intratumoral lymphoid structures. J Clin Oncol 2008;26:4410-7.

10. Chockley PJ, Chen J, Chen G, et al. Epithelial-mesenchymal transition leads to NK cell-mediated metastasis-specific immunosurveillance in lung cancer. J Clin Invest 2018;128:1384-96.

11. Mami-Chouaib F, Blanc C, Corgnac S, et al. Resident memory T cells, critical components in tumor immunology. J Immunother Cancer 2018;6:87.

12. Saigi M, Alburquerque-Bejar JJ, Mc Leer-Florin A, et al. MET-Oncogenic and JAK2-Inactivating Alterations Are Independent Factors That Affect Regulation of PD-L1 Expression in Lung Cancer. Clin Cancer Res 2018;24:4579-87.

13. Cassetta L, Pollard JW. Cancer immunosurveillance: role of patrolling monocytes. Cell Res 2016;26:3-4.

14. Steven A, Fisher SA, Robinson BW. Immunotherapy for lung cancer. Respirology 2016;21:821-33.

15. Reck M, Heigener D, Reinmuth N. Immunotherapy for small-cell lung cancer: emerging evidence. Future Oncol 2016;12:931-43.

16. Villanueva N, Bazhenova L. New strategies in immunotherapy for lung cancer: beyond PD-1/PD-L1. Ther Adv Respir Dis 2018;12:1753466618794133.

17. Gajewski TF, Schreiber H, Fu YX. Innate and adaptive immune cells in the tumor microenvironment. Nat Immunol 2013;14:1014-22.

18. Yang Y. Cancer immunotherapy: harnessing the immune system to battle cancer. J Clin Invest 2015;125:3335-7.

19. Xu F, Zhang H, Chen J, et al. Immune signature of T follicular helper cells predicts clinical prognostic and therapeutic impact in lung squamous cell carcinoma. Int Immunopharmacol 2020;81:105932.

20. Seo JS, Lee JW, Kim A, et al. Whole Exome and Transcriptome Analyses Integrated with Microenvironmental Immune Signatures of Lung Squamous Cell Carcinoma. Cancer Immunol Res 2018;6:848-59.

21. Angarica VE, Del Sol A. Bioinformatics Tools for Genome-Wide Epigenetic Research. Adv Exp Med Biol 2017;978:489-512.

22. Mulder NJ, Adebiyi E, Adebiyi M, et al. Development of Bioinformatics Infrastructure for Genomics Research. Glob Heart 2017;12:91-8.

23. Chen B, Khodadoust MS, Liu CL, et al. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. Methods Mol Biol 2018;1711:243-59.

24. Hu WS, Hsieh MH, Lin CL. A novel atrial fibrillation prediction model for Chinese subjects: a nationwide cohort investigation of 682 237 study participants with random forest model. Europace 2019;21:1307-12.

25. Lim S, Phillips JB, Madeira da Silva L, et al. Interplay between Immune Checkpoint Proteins and Cellular Metabolism. Cancer Res 2017;77:1245-9.

26. Budczies J, Klauschen F, Sinn BV, et al. Cutoff Finder: a comprehensive and straightforward Web application enabling rapid biomarker cutoff optimization. PLoS One 2012;7:e51862.

27. Liu X, Wu S, Yang Y, et al. The prognostic landscape of tumor-infiltrating immune cell and immunomodulators in lung cancer. Biomed Pharmacother 2017;95:55-61.

28. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144:646-74.

29. Qiu L, Yu Q, Zhou Y, et al. Functionally impaired follicular helper T cells induce regulatory B cells and CD14(+) human leukocyte antigen-DR(-) cell differentiation in non-small cell lung cancer. Cancer Sci 2018;109:3751-61.

30. Shang B, Liu Y, Jiang SJ, et al. Prognostic value of tumor-infiltrating FoxP3+ regulatory T cells in cancers: a systematic review and meta-analysis. Sci Rep 2015;5:15179.

31. Liu X, Yu H, Yan C, et al. Plasmacytoid Dendritic Cells and ICOS(+) Regulatory T Cells Predict Poor Prognosis in Gastric Cancer: A Pilot Study. J Cancer 2019;10:6711-5.

32. Zhang Z, Ma L, Goswami S, et al. Landscape of infiltrating B cells and their clinical significance in human hepatocellular carcinoma. Oncoimmunology 2019;8:e1571388.

33. de Mey S, Jiang H, Wang H, et al: Potential of memory T cells in bridging preoperative chemoradiation and immunotherapy in rectal cancer. Radiother Oncol 2018;127:361-9.

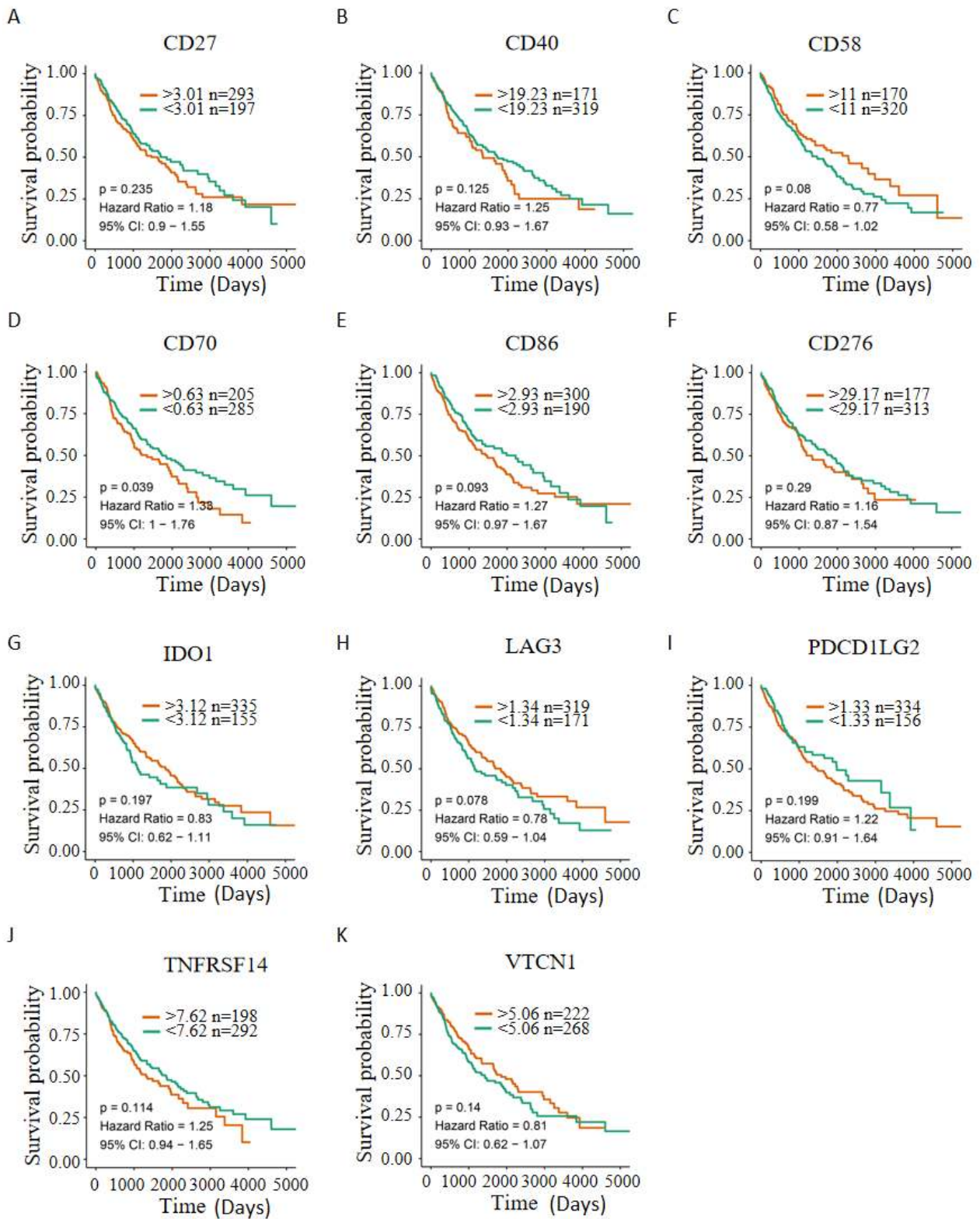**Figure S1** (A-K) Kaplan-Meier curve of different immune modulators in LUSC. Hazard ratio with 95% CI and log rank P value were calculated. P<0.05 was considered significant.
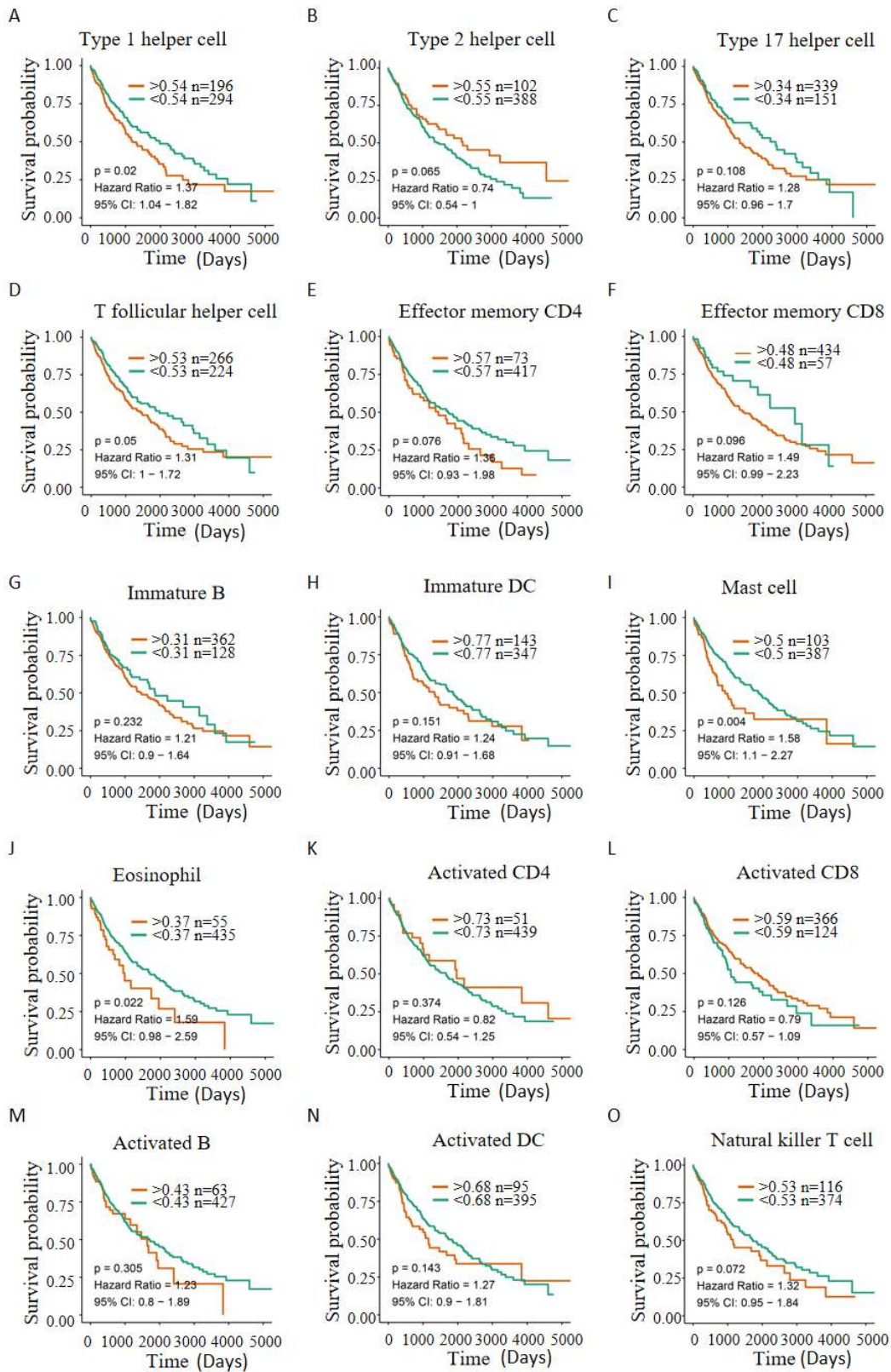
Figure S2 (A-O) Kaplan-Meier curve of TIC distribution patterns in LUSC. Hazard ratio with 95% CI and log rank P value were calculated. P<0.05 was considered significant.